Formulas for Computing Infrastructures

(written by Davide Giannubilo)

How to calculate $T_{I/O}$ for an HDD given the number of blocks

$$T_{I/O} = [(1 - DL) * (T_{seek} + T_{latency}) + T_{transfer} + T_{CTRL}] * N$$

$$T_{latency} = \frac{T_{rotation}}{2} \rightarrow T_{rotation} = \frac{60 * 1000}{RPM}$$

If I don't know RPM (o T_{seek}) → $N * (1 - DL) * T_{service\ time} + N * DL * (T_{transfer} + T_{CTRL})$

RAID Disks: Reliability Calculation

MTTF = Mean Time To Failure

- RAID 0
  - $MTTF_{RAID\ 0} = MTTF_{single\ disk}/N$
- RAID 1 (per N=2)
  - $MTTF_{RAID\ 1} = (MTTF_{single\ disk}/N) * (1/(MTTR/MTTF_{single\ disk})$
- RAID 1+0
  - $MTTF_{RAID\ 1+0} = (MTTF_{single\ disk}{}^2)/(N * MTTR)$
- RAID 0+1
  - $MTTF_{RAID\ 0+1} = (MTTF_{single\ disk}{}^2)/(N * G * MTTR)$
    - $G = (N/2)$ is the number of disks in a stripe group
- RAID 4
  - $MTTF_{RAID\ 4} = (MTTF_{single\ disk}{}^2)/(N * N - 1 * MTTR)$
- RAID 5
  - $MTTF_{RAID\ 5} = (MTTF_{single\ disk}{}^2)/(N * N - 1 * MTTR)$
- RAID 6
  - $MTTF_{RAID\ 6} = 2 * (MTTF_{single\ disk}{}^3)/(N * N - 1 * N - 2 * MTTR)$

	RAID 0	**RAID 1**	**RAID 4**	**RAID 5**	**RAID 6**
Capacity	N	N / 2	N – 1	N – 1	N – 2
Seq. Read	N * S	(N / 2) * S	(N – 1) * S	(N – 1) * S	
Seq. Write	N * S	(N / 2) * S	(N – 1) * S	(N – 1) * S	
Random Read	N * R	N * R	(N – 1) * R	N * R	
Random Write	N * R	(N / 2) * R	R / 2	(N / 4) * R	
Reliability (how many disk(s) can fail)	NO	1 (or N / 2)	1	1	2

Best performance and most capacity? → RAID 0
Greatest error recovery? → RAID 1 (1+0 or 0+1) or RAID 6
Balance between space, performance, and recoverability? → RAID 5

Performance

Operational laws are based on observable variables:

- T, the length of time we observe the system
- A_k, the number of request arrivals we observe for resource k
- C_k, the number of request completions we observe at resource k
- B_k, the total amount of time during which the system is busy
- N_k, the average number of jobs in the resource k

From these observed values, we can derive the following 4 quantities:

1. $\lambda_k = A_k/T$, the arrival rate
2. $X_k = C_k/T$, the throughput or completion rate
3. $U_k = B_k/T$, the utilisation
4. $S_k = B_k/C_k$, the mean service time per completed job

(The upper formulas are also valid if we have total data and not partial ones like $X = C/T$)

Utilization Law: $U_k = X_k * S_k \; and \; X_k = U_k/S_k$
Little's Law: $N_{insystem} = X * R$
Response Time Law: $R = N/X - Z$ (where N number of users, X system throughput and Z think time)
Forced Law: $X_k = v_k * X$

Other laws:

$V_k = C_k/C$	$R_k(residence\ time) = V_k * R_k\ (response\ time)$
$D_k = V_k * S_k = B_k/C$	$R = \sum R_k$
$U_k = X * D_k$	$N = \sum N_k$
$X_{sys} = U_k/D_k$	

$R = \sum_{i=1}^{k} \frac{D_k}{1 - \lambda D_k}$ $\qquad$ $R(\lambda) = \frac{D_{CPU}}{1 - \lambda\, D_{CPU}} + \frac{D_{disk}}{1 - \frac{\lambda\, D_{disk}}{n}}$

Fundamentals laws:

$D = \sum_{k=1}^{K} D_k$ and $D_{max} = \max_k D_k$ $\qquad$ $N^* = \frac{D+Z}{D_{max}}$

batch	$\frac{1}{D} \le X(N) \le \min\left(\frac{N}{D}, \frac{1}{D_{max}}\right)$
terminal	$\frac{N}{ND+Z} \le X(N) \le \min\left(\frac{N}{D+Z}, \frac{1}{D_{max}}\right)$
transaction	$X(\lambda) \le \frac{1}{D_{max}}$

batch	$\max(D, ND_{max}) \le R(N) \le ND$
terminal	$\max(D, ND_{max} - Z) \le R(N) \le ND$
transaction	$D \le R(\lambda)$

Transaction = Open Models
Batch & Terminal = Closed Models

Dependability

MTTF: mean time to (first) failure, the up time before the first failure

MTBF: mean time between failures $= \frac{total\ operating\ time}{number\ of\ failures}\ or\ \frac{1}{\lambda}\ where\ \lambda = \frac{number\ of\ failures}{total\ operating\ time}$

Block diagram:

- **Component in series**
    a. 2 blocks, C1 and C2

$$R_S = R_{C1} * R_{C2}$$

In general, if system S is composed by components with a reliability having an exponential distribution (very common case)

$$R_s(t) = e^{-\lambda_s t}\ where \lambda_s = \sum_{i=1}^{n} \lambda_i \rightarrow MTTF_s = \frac{1}{\sum 1/MTTF_i}$$

A special case: when all components are identical

$$R_s(t) = e^{-\lambda_s t}$$

$$R_S(t) = e^{-n\lambda t} = e^{-\frac{nt}{MTTF_1}} \qquad MTTF_S = \frac{MTTF_1}{n}$$

Availability:

$$A_S = \prod_{i=1}^{n} \frac{MTTF_i}{MTTF_i + MTTR_i}$$

When all components are the same:

$$A_S(t) = A_1(t)^n \qquad A = \left(\frac{MTTF_1}{MTTF_1 + MTTR_1} \right)^n$$

- **Component in parallel**
    a. 2 blocks, C1 and C2

$$R_S = R_{C1} + R_{C2} - R_{C1} * R_{C2}$$

$MTTF_{parallel} = MTTF * \sum_{n=1}^{N} \frac{1}{n}$ (if all MTTF are equal)

$MTTF_{parallel} = \sum_{i=1}^{N} MTTF_i - \frac{1}{\sum_{i=1}^{N} \frac{1}{MTTF_i}}$ (if all MTTF are different)

System P composed by *n* components

$$R_P(t) = 1 - \prod_{i=1}^{n} \left(1 - R_i(t)\right)$$

Availability

$$A_P(t) = 1 - \prod_{i=1}^{n} \left(1 - A_i(t)\right)$$

$$A_P = 1 - \prod_{i=1}^{n} \left(1 - A_i\right) = 1 - \prod_{i=1}^{n} \frac{MTTR_i}{MTTF_i + MTTR_i}$$